

УДК 004

Генаш Максим Геннадійович, Олійник Володимир Валентинович
Національний технічний університет України "Київський політехнічний
інститут імені Ігоря Сікорського"
(Київ, Україна)

ПІДХІД ДО РЕАЛІЗАЦІЇ СИСТЕМИ МОДЕЛЮВАННЯ ЗОВНІШНОСТІ У РЕАЛЬНОМУ ЧАСІ НА ОСНОВІ СЕМАНТИЧНОЇ СЕГМЕНТАЦІЇ ОБЛИЧЧЯ

Анотація. У даній роботі описано підхід до реалізації WEB-додатку для віртуальної зміни зачіски та кольору волосся у потоці відео з камери пристрою (ноутбука, смартфона, тощо) за допомогою навченої моделі (заснованої на нейронних мережах) для семантичної сегментації обличчя. На відміну від існуючих проектів, у даній роботі запропоновано підхід заснований на семантичній сегментації, а не пошуку лиш ключових точок обличчя, що дозволяє якісно підставляти нову зачіску навіть якщо вона коротша за наявну зараз. Крім того заміна відбувається прямо у потоці відео з камери, а не на статичних фотографіях, що стало можливим завдяки використанні навченої моделі прямо в браузері без необхідності обміну даними з сервером для їх обробки.

Ключові слова: сегментація, нейронні мережі, Tensorflow.js, PWA, SPA, доповнена реальність, обробка фотографій, обробка обличчя, зміна зачіски.

Генаш Максим Геннадьевич, Олейник Владимир Валентинович
Национальный технический университет Украины
"Киевский политехнический институт имени Игоря Сикорского"
(Киев, Украина)

ПОДХОД К РЕАЛИЗАЦИИ СИСТЕМЫ МОДЕЛИРОВАНИЯ ВНЕШНОСТИ В РЕАЛЬНОМ ВРЕМЕНИ НА ОСНОВЕ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ ЛИЦА

Аннотация. В данной работе описан подход к реализации WEB-приложения по виртуальной смене причёски и цвета волос в потоке видео с камеры устройства (ноутбука, смартфона и т.д.) с помощью обученной модели (основанной на нейронных сетях) для семантической сегментации лица. В отличие от существующих проектов в данной работе предложен подход основан на семантической сегментации, а не поиска только ключевых точек лица, что позволяет качественно подставлять новую причёску даже если она короче имеющейся сейчас. Кроме того замена происходит прямо в потоке видео с камеры, а не на статических фотографиях, что стало возможным благодаря использованию обученной модели прямо в браузере без необходимости обмена данными с сервером для их обработки.

Ключевые слова: сегментация, нейронные сети, Tensorflow.js, PWA, SPA, дополненная реальность, обработка фотографий, обработка лица, смена причёски.

*Maksym Henash; Ph.D., assistant professor Volodymyr Oliinyk
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
(Kyiv, Ukraine)*

APPROACH TO THE IMPLEMENTATION OF THE SYSTEM FOR APPEARANCE MODELING IN REAL TIME BASED ON SEMANTIC FACE SEGMENTATION

Abstract. *This paper describes research on ability and feasibility of applying neural networks of UNet, DeepLabV3, PSPNet architectures in semantic segmentation of faces. The training was performed on Labeled Faces in the Wild (LFW) Part Labels Database. Semantic segmentation was performed by 3 classes: hair, face region, background. As the result of the research it was achieved fairly high level of segmentation accuracy for model UNet (Mean IoU = 85.6%, Pixel Accuracy = 95.7%) which is comparable with results of state of the art models on LFW dataset, meanwhile the trained model is compact enough to be appropriate for using in mobile and web applications. This paper describes an approach to the implementation of a WEB application for virtual changing of hairstyles and hair colors in a video stream from a device camera (laptop, smartphone, etc.) using a trained model (based on neural networks) for semantic segmentation of a face. Unlike existing projects, this paper proposes an approach based on semantic segmentation, rather than finding only the key points of the face, which allows you to substitute a new hairstyle even if it is shorter than the current one. In addition, the replacement takes place directly in the video stream from the camera, and not in static photos, which became possible by using of the trained model directly in the browser without the need to exchange data with the server to process them.*

Keywords: *segmentation, neural networks, Tensorflow.js, PWA, SPA, augmented reality, photo processing, face editing, hairstyle editing.*

1. Вступ

Останні роки завдяки росту потужності комп'ютерів та мобільних пристроїв і розвитку сфери штучного інтелекту (зокрема, в даному випадку, комп'ютерного зору) почала активно розвиватися і привертати увагу (як розробників так і користувачів) індустрія доповненої реальності.

У відповідності до [1], доповнена реальність (англ. augmented reality, AR), — термін, що позначає всі проекти, спрямовані на доповнення реальності будь-якими віртуальними елементами.

Багато людей стикалися хоч раз у житті із проблемою: вони хочуть спробувати змінити у своїй зовнішності якісь риси, наприклад зачіску, чи колір волосся, але не впевнені чи це пасуватиме їм.

Завдяки розвитку систем доповненої реальності тепер можна провести зміну у своїй зовнішності віртуально.

На даний момент існує ряд програм, за допомогою яких користувач може сам перевірити чи пасуватиме йому якась зачіска / стрижка, колір волосся, тощо, проте ці програми або занадто обмежені, або вимагають забагато зусиль та спеціалізованих знань.

У даній роботі запропоновано одне із можливих рішень, що базується на найсучасніших технологіях і враховує недоліки існуючих проектів.

2. Існуючі підходи

Можна виділити декілька існуючих підходів до вирішення віртуальної заміни зачіски:

1) Повністю мануальна зміна за допомогою професійних фото редакторів, таких як, наприклад, Adobe Photoshop [2] – найбільш гнучкий спосіб, але вимагає суттєвих часових затрат (від пів-години до декількох годин в залежності від рівня володіння інструментом та перфекціонізму користувача), наявність встановлених на комп'ютері фото редакторів (ліцензія на які зазвичай коштує дорожче за десяток зачісок) і високий рівень володіння цими фото редакторами (освоєння яких є абсолютно неоправданими часовими затратами для більшості пересічних користувачів).

2) Мануальна підстановка за допомогою спеціалізованих програм (наприклад [3], [4], [5]) у режимі drag and drop – значно швидший (від декількох секунд до хвилини) і простіший спосіб, який не вимагає спеціалізованих знань, проте обмежений набором наявних у програмі списком зачісок і дуже не точний, адже нова зачіска накладається поверх існуючої, що робить неможливим перевірку чи пасуватиме людині коротша стрижка, не кажучи вже про врахування кута повороту голови та рівня освітленості фотографії.

3) Автоматизована підстановка (наприклад [6]): користувач вручну виділяє область обличчя, а програма вже сама підставляє навколо цієї області нову зачіску – спосіб нічим не кращий за попередній, має такі ж недоліки.

4) Автоматична підстановка нової зачіски на фотографію за допомогою автоматичного пошуку ключових точок на обличчі такими методами як AAM (Active Appearance Model), ASM (Active Shape Model), CLM (Constrained Local Models), тощо (наприклад [7], [8]) – суттєво зручніший та швидший спосіб за попередні, можлива підстановка нових зачісок прямо у потоці відео, можливе врахування кута повороту голови, проте все ще наявний недолік із неможливістю підстановки на фотографію коротшої стрижки через те, що програма не знає яку область охоплює волосся на наявній фотографії.

Недоліки описаних методів можна виправити завдяки застосуванню системи семантичної сегментації обличчя на фотографії.

3. Запропонований підхід

3.1. Принцип роботи

На даний момент WEB-браузер є універсальною платформою, яка дозволяє запускати розроблені додатки на різних пристроях (комп'ютери, ігрові консолі, шоломи віртуальної реальності, смартфони, телевізори, smart-watch) та різних операційних системах (ОС), роблячи розроблену програму доступною для користування половині населення світу (усім в кого є доступ до мережі Інтернет). Саме тому у цій роботі описано підхід до створення саме WEB-додатку, адже це є перспективнішою платформою.

Далі описано принцип роботи запропонованої системи.

Система являється Progressive Web Application (PWA) [9]. При першому відкритті користувачем сайту з додатком він завантажується із сервера у кеш браузера і при наступних відкриттях завантажується із кеша. Раз на добу веб-додаток у фоновому режимі відправляє запит на сервер для перевірки наявності нової версії.

Після запуску додатку користувачеві пропонується обрати зачіску та її колір із підготовленого списку. Зачіски являються PNG зображеннями із

прозорим фоном. Кожне зображення формується на основі обробленої фотографії зробленої у анфас реальної зачіски людини із вирізаним фоном та підготовленим набором кольорів.

Після вибору зачіски та її кольору користувач може підставити її на завантажену фотографію чи у потік відео з камери. У випадку підстановки зачіски на відео окремі кадри обробляються як незалежні зображення, при цьому після обробки одного кадра береться не наступний у черзі, а той що надійшов останнім на даний момент; це не обхідно для врахування обмежень у швидкості роботи моделі семантичної сегментації, яка в залежності від реалізації й потужності пристрою на якому запущено додаток іноді може виконувати обробку довше ніж 33мс (стільки необхідно для забезпечення 30кадрів на секунду).

Семантична сегментація виконується з використанням бібліотеки Tensorflow.js [10], який дозволяє запускати навчені моделі нейронних мереж прямо в браузері. При цьому навчання можна проводити за допомогою бібліотек Tensorflow, Keras у програмах написаних на Python чи C++ на потужному обладнанні з використанням GPU/TPU і далі просто конвертувати [11] навчені моделі у формат необхідний для Tensorflow.js.

Варто також зауважити, що модель для семантичної сегментації має бути достатньо компактною (бажано не більше 100Мб, в ідеалі – до 10Мб), адже при першому запуску додатку вона буде завантажуватися із сервера в браузер, тож при великих розмірах моделі це може зайняти багато часу.

3.2. Алгоритм заміни зачіски

Після вибору користувачем бажаної зачіски та завантаження його фотографії застосовується наступний алгоритм:

1) Копія фотографії масштабується до розміру 512x512px (такий розмір є найбільш вдалим компромісом між високою якістю фотографії й швидкодією системи) і подається на вхід моделі семантичної сегментації. На виході з моделі отримується зображення 512x512px, на якому сегментовано піксели, що відповідають 4 класам: волосся, контур обличчя, очі та фон (усе інше).

2) На сегментованому зображенні вирізається область, що відповідає волосся.

3) Контур обличчя у вирізаних областях доповнюється середнім кольором таким чином, щоб отримати зображення усієї голови без волосся із заокругленим верхом. Верхня точка доповненого обличчя відповідає відстані, яка рівна 7 відстаням між крайньою лівою точкою правого ока та крайньою правою точкою лівого ока. Якщо обрана зачіска була «лиса голова» - переходимо до пункту 5).

4) На доповнене обличчя підставляється масштабоване зображення зачіски. Для кожної зачіски попередньо розраховані пропорції відстаней на яких треба розмістити зачіску відносно області очей.

5) Оброблене зображення масштабується до розміру вихідної фотографії і показується користувачу.

При використанні потоку відео з камери застосовується той же алгоритм, але для кожного кадру окремо.

4. Висновок

В результаті проведеного дослідження у даній роботі запропоновано підхід до реалізації WEB-додатку для віртуальної зміни зачіски та кольору

волосся на фотографії, чи у потоці відео з камери пристрою за допомогою системи семантичної сегментації обличчя на фотографії.

Запропонована система враховує недоліки існуючих проектів, являючись зручнішою для використання і надаючи можливість зміни зачісок навіть у випадках коли нова зачіска є коротшою за попередню.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ТА ЛІТЕРАТУРИ:

1. Доповнена реальність [Електронний ресурс] // Вікіпедія. – 2018. – Режим доступу до ресурсу: https://uk.wikipedia.org/wiki/Доповнена_реальність.
2. Adobe Photoshop CC [Електронний ресурс] // Adobe. – 2018. – Режим доступу до ресурсу: <https://www.adobe.com/ua/products/photoshop.html>.
3. Man HairStyle Photo Editor 2018 [Електронний ресурс] // Google Play – Режим доступу до ресурсу: <https://play.google.com/store/apps/details?id=com.bebal.menhair.changer>.
4. Man HairStyle Photo Editor [Електронний ресурс] // Google Play. – 2018. – Режим доступу до ресурсу: <https://play.google.com/store/apps/details?id=man.hair.manhairstylephoto>.
5. Virtual Makeover [Електронний ресурс] // Google Play. – 2018. – Режим доступу до ресурсу: <https://play.google.com/store/apps/details?id=com.modiface.virtualmakeover>.
6. Woman & Girl Hair Styler App - Hair Color Changer [Електронний ресурс] // Google Play. – 2018. – Режим доступу до ресурсу: <https://play.google.com/store/apps/details?id=com.greenzero.hairtrendx>.
7. YouCam Makeup [Електронний ресурс] // Google Play. – 2018. – Режим доступу до ресурсу: <https://play.google.com/store/apps/details?id=com.cyberlink.youcammakeup>.
8. MSQRD [Електронний ресурс] // Google Play. – 2018. – Режим доступу до ресурсу: <https://play.google.com/store/apps/details?id=me.msqrd.android>.
9. Progressive Web Apps [Електронний ресурс] // Google Developers. – 2018. – Режим доступу до ресурсу: <https://developers.google.com/web/progressive-web-apps/>.
10. TensorFlow.js [Електронний ресурс] // TensorFlow. – 2018. – Режим доступу до ресурсу: <https://js.tensorflow.org/>.
11. Importing a Keras model into TensorFlow.js [Електронний ресурс] // TensorFlow. – 2018. – Режим доступу до ресурсу: <https://js.tensorflow.org/tutorials/import-keras.html>.